

Optical Diffraction in Close Proximity to Plane Apertures. I. Boundary-Value Solutions for Circular Apertures and Slits

Volume 107

Number 4

July–August 2002

Klaus D. Mielenz

National Institute of Standards and Technology,
Gaithersburg, MD 20899-8440

meezlink@gcnetmail.net

In this paper the classical Rayleigh-Sommerfeld and Kirchhoff boundary-value diffraction integrals are solved in closed form for circular apertures and slits illuminated by normally incident plane waves. The mathematical expressions obtained involve no simplifying approximations and are free of singularities, except in the aperture plane itself. Their use for numerical computations was straightforward and provided new insight into the nature of diffraction in the near zone where the Fresnel approximation does not apply. The Rayleigh-Sommerfeld integrals were found to be very similar to each other, so that polarization effects appear to be

negligibly small. On the other hand, they differ substantially at sub-wavelength differences from the aperture plane and do not correctly describe the diffracted field as an analytical continuation of the incident geometrical field.

Key words: boundary value theory; circular apertures; diffraction; Kirchhoff; near zone; optics; polarization; Rayleigh; Sommerfeld; scalar wave functions; slits.

Accepted: July 29, 2002

Available online: <http://www.nist.gov/jres>

1. Introduction

Diffraction problems in optics typically involve distances from the diffracting screen which are large in comparison to the wavelength of light. Accordingly the Fresnel and Fraunhofer approximations of the principal classical diffraction integrals are well documented, but so far no workable expressions have been available for computations in the near zone. The aim of the present paper is to develop mathematical procedures for the latter purpose and use them to study the behavior of these integrals in the proximity of plane apertures.

We begin by citing the classical scalar expressions for analyzing optical diffraction by an aperture; namely, Kirchhoff's integral equation [1]

$$U(P) = -\frac{1}{4\pi} \int_{\mathcal{S}} dQ \left[U(Q) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ikQP}}{QP} \right) - \frac{\partial U(Q)}{\partial \mathbf{n}} \frac{e^{ikQP}}{QP} \right] \quad (1)$$

and, alternatively, the Rayleigh-Sommerfeld integral equations [2, 3, 4]

$$\begin{aligned} U(P) &= -\frac{1}{2\pi} \int_{\mathcal{S}} dQ \frac{\partial U(Q)}{\partial \mathbf{n}} \frac{e^{ikQP}}{QP} \\ &= \frac{1}{2\pi} \int_{\mathcal{S}} dQ U(Q) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ikQP}}{QP} \right), \end{aligned} \quad (2)$$

In these equations, U is a scalar wave function, \mathcal{S} is a closed surface containing a plane aperture \mathcal{A} located in the xy -plane of a cartesian coordinate system as indicated in Fig. 1, $P = (x, y, z)$ is the point of observation ($z \geq 0$), $Q = (\xi, \eta, 0)$ is a point on \mathcal{A} , QP is the distance between them, \mathbf{n} is the aperture normal pointing in the direction of the positive z -axis, and $k = 2\pi/\lambda$ is the circular wavenumber of monochromatic light with wavelength λ .

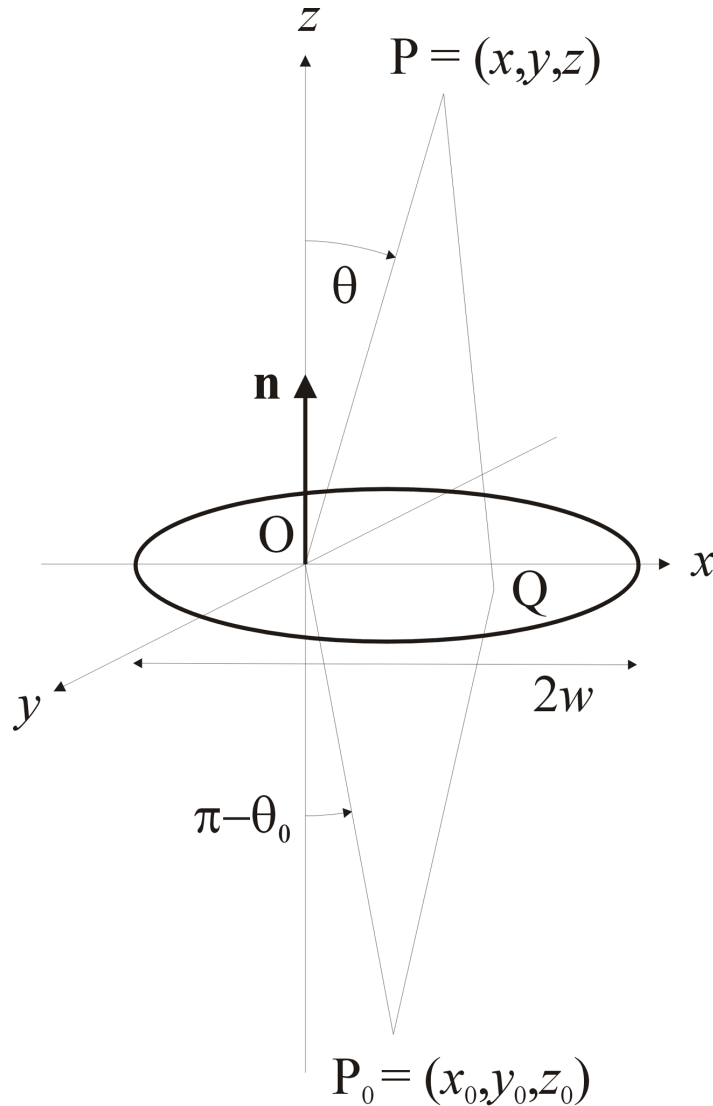


Fig. 1. Basic geometry for plane aperture of width $2w$. \mathbf{n} = aperture normal, O = coordinate origin, P_0 = point source, P = point of observation.

2. Background

The Kirchhoff and Rayleigh-Sommerfeld integral equations (1) and (2) are alternative forms of the theorem of Helmholtz [5], which expresses Huygens' principle in terms of a scalar wave function U and its normal derivatives without assuming specific attributes of this function, except that it is continuous and twice differentiable with continuous derivatives and obeys the homogeneous wave equation,

$$\Delta U + k^2 U = 0, \quad (3)$$

on and within the closed surface \mathcal{S} . As Helmholtz'

theorem by itself is insufficient to provide a unique solution, it is necessary to impose additional constraints on U by prescribing its boundary values on \mathcal{S} .

Kirchhoff considered a "black screen which neither reflects nor transmits light." He assumed, plausibly, that in this case the incident field vanishes altogether on the opaque portion of the screen and is equal to the unperturbed incident field $U_{\text{geom}}(Q)$ inside the aperture. Thus,

$$U(Q) = 0 \text{ and } \frac{\partial U(Q)}{\partial \mathbf{n}} = 0, \text{ when } Q \notin \mathcal{A}, \quad (4a)$$

$$U(Q) = U_{\text{geom}}(Q), \text{ when } Q \in \mathcal{A}, \quad (4b)$$

so that Eq. (1) is reduced to Kirchhoff's familiar formula, where the integration extends over the aperture area \mathcal{A} only, and U is replaced by U_{geom} in the integrand. As it turned out, Kirchhoff's solution is mathematically flawed. Poincaré [6] discovered that it contradicts itself and predicted that it will not reproduce the assumed boundary conditions Eq. (4a,b).

Sommerfeld [4, 7] recognized that these difficulties are due to the fact that U and $\partial U/\partial \mathbf{n}$ cannot both vanish on any finite portion of the closed surface \mathcal{S} unless U is everywhere identically equal to zero. He remedied the problem by deriving the integral equations [Eq. (2)] which require only the boundary values of either U or $\partial U/\partial \mathbf{n}$ to specify a solution. Thus he assumed, instead of Eq. (4a),

$$\frac{\partial U(\mathbf{Q})}{\partial \mathbf{n}} = 0 \text{ or } U(\mathbf{Q}) = 0, \text{ when } \mathbf{Q} \notin \mathcal{A}, \quad (5a)$$

$$U(\mathbf{Q}) = U_{\text{geom}}(\mathbf{Q}), \text{ when } \mathbf{Q} \in \mathcal{A}, \quad (5b)$$

and applied these conditions separately to the first and second Eqs. (2) Except for a passing reference to “*shiny screens*,” Sommerfeld did not explicitly associate his theory with the diffraction of polarized light but merely offered the solution appearing in Eq. (5e), below, as a mathematically improved alternative to Kirchhoff's formula. On the other hand, Rayleigh [2,3] emphasized that the screen must be supposed to be a perfect metallic reflector so that the boundary conditions Eq. (5a) pertain to p - and s -polarized incident light, respectively. He emphasized, further, that these conditions apply on the dark side of the screen only, while on the lit side it must be assumed that

$$U_{\text{geom}}^{(p)}(\mathbf{Q}) = 0 \text{ or } \frac{\partial U_{\text{geom}}^{(s)}(\mathbf{Q})}{\partial \mathbf{n}} = 0, \text{ when } \mathbf{Q} \notin \mathcal{A}. \quad (5c)$$

According to Maxwell's equations, $U^{(p)}$ and $\partial U^{(s)}/\partial \mathbf{n}$ must be continuous at the screen and hence it follows that, taken together, the conditions of Eqs. (5a) and (5c) stipulate that for either state of polarization U and $\partial U/\partial \mathbf{n}$ must both zero on the dark side of the screen. This is the same as Kirchhoff's boundary condition of Eq. (4a), but without mathematical contradictions. The final result for the Rayleigh-Sommerfeld integrals is

$$U_{\text{RS}}^{(p)}(\mathbf{P}) = -\frac{1}{2\pi} \int_{\mathcal{A}} d\mathbf{Q} \frac{\partial U_{\text{geom}}(\mathbf{Q})}{\partial \mathbf{n}} \frac{e^{ikQP}}{QP}, \quad (5d)$$

$$U_{\text{RS}}^{(s)}(\mathbf{P}) = \frac{1}{2\pi} \int_{\mathcal{S}} d\mathbf{Q} U_{\text{geom}}(\mathbf{Q}) \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ikQP}}{QP} \right), \quad (5e)$$

where, as in Kirchhoff's theory, the screen itself does

not contribute to the integrals. According to an analysis performed by Mukunda [8], these expressions do recover the assumed boundary conditions in Eqs. (5a-c) as $\mathbf{P} \rightarrow \mathbf{Q}$, in the sense that $U_{\text{RS}}^{(s)}$ replicates the assumed value of U_{geom} but not necessarily the compatible value of $\partial U_{\text{geom}}/\partial \mathbf{n}$, and the converse is true for $U_{\text{RS}}^{(p)}$.

It should also be noted that Kirchhoff's solution is simply the arithmetic mean of the Rayleigh-Sommerfeld solutions,

$$u_{\text{K}}(\mathbf{P}) = \frac{1}{2} [u_{\text{RS}}^{(p)}(\mathbf{P}) + u_{\text{RS}}^{(s)}(\mathbf{P})], \quad (6a)$$

and that in the so-called Fresnel limit they are all reduced to one and the same expression. For a point source $\mathbf{P}_0 = (x_0, y_0, z_0)$, as shown in Fig. 1, and assuming that the distances $-z_0$ and z are large in comparison to the aperture width $2w$ and the wavelength λ , and that $\cos(\pi - \theta_0)$ and $\cos \theta$ are essentially equal to 1, one finds [8]

$$U_{\text{K}}(\mathbf{P}) \approx U_{\text{RS}}^{(p)}(\mathbf{P}) \approx U_{\text{RS}}^{(s)}(\mathbf{P}) \approx U_{\text{F}}(\mathbf{P}) = \frac{i\sqrt{I_0}e^{ik(z-z_0)}}{zz_0} \int_{\mathcal{A}} d\mathbf{Q} e^{ik\Delta(\mathbf{Q})}, \quad (6b)$$

where I_0 is the radiant intensity of the source and $\Delta(\mathbf{Q})$ is a second-order approximation of the path difference $(P_0Q + QP) - (P_0O - OP)$. It may be estimated that this approximation is accurate to 1 % or better when $z > 20w$ and $z > 20\lambda$ so that it is usually satisfied for narrow apertures and short wavelengths; say, $2w = 0.1$ mm and $\lambda = 500$ nm, as used for pinhole imagery or classroom experiments. In these cases the Rayleigh-Sommerfeld and Kirchhoff solutions will hardly be needed in their rigorous forms, but on the other hand the reliability of the Fresnel approximation is doubtful for large apertures and wavelengths. It may not be applicable in the focal planes of fast lenses or in the case of large apertures used in radiometry and photometry, especially in the infrared and microwave regions.

Apart from the above, it appears that the behavior of the Kirchhoff and Rayleigh-Sommerfeld integrals in the aperture plane has not been documented in the literature except in two isolated cases. Wolf and Marchand [10] derived a closed expression for Kirchhoff's integral U_{K} inside a circular aperture illuminated by a normally incident plane wave, using the Maggi-Rubinowicz transformation [11] of U_{K} and a stationary-phase approximation. As shown in Fig. 2, this expression gives a fair indication of an oscillatory behavior of U_{K} in the aperture plane but exhibits spurious singularities at the aperture center and rim. Osterberg and Smith [12] found a closed expression for $U_{\text{RS}}^{(s)}$ at axial points of observation

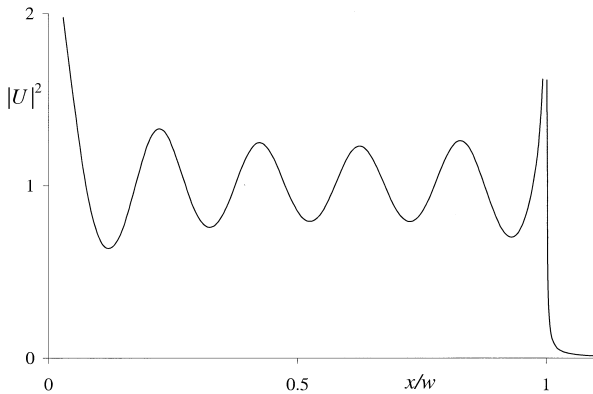


Fig. 2. Approximate behavior of Kirchhoff's integral U_K inside a circular aperture of diameter $2w = 10\lambda$. Computed from Eq. (A16) of Ref. [10].

behind a circular aperture and confirmed that it does represent a continuous extension of the incident field into the half space $z > 0$. On the other hand, it is evident from Fig. 3 that the combined field is not continuously differentiable at $z = 0$, showing that $U_{RS}^{(s)}$ still does not fully meet the requirements of Helmholtz' theorem. There appear to be no known solutions for $U_{RS}^{(p)}$, but based on Mukunda's work it is clear that $U_{RS}^{(p)} = 2U_K - U_{RS}^{(s)}$ is discontinuous in the aperture plane because U_K is discontinuous.

In addition to these publications, the pertinent literature also contains a considerable number of papers that were, for the most part, intended to "save" Kirchhoff's theory in one way or another. For example, Kottler [13] regarded Kirchhoff's integral as the rigorous solution of a "saltus" problem in which the wave function U has prescribed discontinuities rather than boundary values at

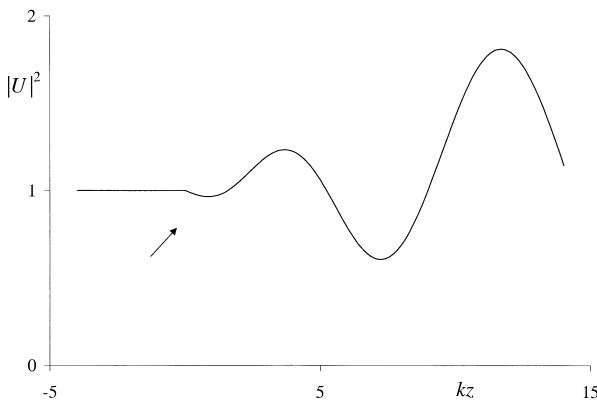


Fig. 3. Behavior of the Rayleigh-Sommerfeld integral $U_{RS}^{(s)}$ for axial points near a circular aperture of diameter $2w = 10\lambda$. Computed from Eq. (7) of Ref. [12].

the aperture screen.¹ Kottler's theory involved the above-mentioned Maggi-Rubinowicz transformation, and subsequently the singularities inherent in the latter led to the revival of a belief that diffraction can be attributed to "boundary diffraction waves" emerging from the edges of aperture screens. Marchand and Wolf [14, 15] asserted that the inconsistencies of Kirchhoff's integral are only apparent and developed a theory in which U_K is expressed in terms of vector potentials that have singularities even in free space. Born [16] suggested the possibility that Kirchhoff's formula may be only one in a series of successive approximations, and Franz [17] re-derived it by an iterative method in which the discontinuities of previous solutions are regarded as secondary sources of light. All in all, this curious exchange of conjectures has raised more questions than it has answered. Most certainly, it has not addressed the concerns of laboratory physicists in search of a "best" theory for practical application.

3. Mathematical Expressions and Numerical Results

3.1 General

In order to analyze the behavior of the Kirchhoff and Rayleigh-Sommerfeld integrals in the proximity of apertures it is necessary to derive usable expressions for computations in the near zone. For this purpose and to keep the calculations simple, it will be assumed in the following that the incident field is a normally incident plane wave so that the geometrical field in the aperture is given by

$$U_{\text{geom}}(\mathbf{Q}) = \sqrt{E_0}, \quad \frac{\partial U_{\text{geom}}(\mathbf{Q})}{\partial \mathbf{n}} = ik\sqrt{E_0}, \quad (7a)$$

where E_0 denotes irradiance, and Eqs. (5d) and (5e) can be written in normalized form as

$$u_{RS}^{(p)}(\mathbf{P}) \equiv U_{RS}^{(p)}(\mathbf{P})/\sqrt{E_0} = -\frac{ik}{2\pi} \int_{\mathcal{A}} d\mathbf{Q} \frac{e^{ikQP}}{QP}, \quad (7b)$$

$$u_{RS}^{(s)}(\mathbf{P}) \equiv U_{RS}^{(s)}(\mathbf{P})/\sqrt{E_0} = \frac{1}{2\pi} \int_{\mathcal{A}} d\mathbf{Q} \frac{\partial}{\partial \mathbf{n}} \left(\frac{e^{ikQP}}{QP} \right) = \frac{1}{ik} \frac{\partial u_{RS}^{(p)}}{\partial z}. \quad (7c)$$

¹ *Saltus* is Latin for *jump* or *leap*. This author prefers the maxim that *natura non facit saltus*.

the limits of integration being from 0 to $w + x$ when $x \leq w$ and from $x - w$ to $x + w$ when $x \geq w$.

For the special case of axial points of observation ($x = 0$) the angle χ defined by Eq. (8c) is zero, so that Eq. (8d) can be solved in closed form. On substitution of $i\beta$ as a new integration variable and subsequent differentiation with respect to z , one finds

$$u_{\text{RS}}^{(p)}(0, z) = e^{ikz} - e^{ikW}, \quad W = \sqrt{w^2 + z^2}, \quad (9a)$$

$$u_{\text{RS}}^{(s)}(0, z) = e^{ikz} - \frac{ze^{ikW}}{W}, \quad (9b)$$

the latter being identical to the above-mentioned expression derived by Osterberg and Smith [13].

For $x \neq 0$, the numerical integration methods described in Ref. [18] were used to find the real and imaginary parts of Eqs. (8d) and (8e) for a small circular aperture of diameter $2w = 10\lambda$ at the distances $z = 0.01\lambda$, λ , and 10λ . The results obtained are shown in Figs. 5a through 5c and will be discussed in Sec. 4.

3.3 Slit

Next we consider a diffracting slit of width $2w$, centered in the xy -plane of a rectangular coordinate system as indicated in Fig. 6. The corresponding diffraction pattern will consist of straight bands which are parallel to the slit jaws, and thus it will again be sufficient to compute its variation in the xz -plane. For a given point of observation $P = (x, 0, z)$ and for arbitrary aperture points $Q = (\xi, \eta, 0)$, Eq. (7b) becomes

$$u_{\text{RS}}^{(p)}(x, z) = -\frac{ik}{2\pi} \int_{-w}^w d\xi \int_{-\infty}^{\infty} d(k\eta) \frac{e^{ik\sqrt{(\xi-x)^2 + \eta^2 + z^2}}}{k\sqrt{(\xi-x)^2 + \eta^2 + z^2}}$$

$$= \frac{k}{2} \int_{-w-x}^{w-x} d(\xi-x) H_0^{(1)}(\beta), \quad \beta = k\sqrt{(\xi-x)^2 + z^2}, \quad (10a)$$

where $H_0^{(1)} = J_0 + iY_0$ is the Hankel function of the first kind and zero order [19, 20], J_0 and Y_0 are the corresponding Bessel functions of the first and second kind, and β is the same as in Eq. (8a). Hence the solution for $u_{\text{RS}}^{(s)}$ is obtained at once by substitution of $\partial H_0^{(1)}/\partial z = -k^2 Z H_1^{(1)}/\beta$ into Eq. (10a), leading to

$$u_{\text{RS}}^{(s)}(x, z) = \frac{1}{ik} \frac{\partial u_{\text{RS}}^{(p)}}{\partial z} = \frac{ik^2 z}{2} \int_{-w-x}^{w-x} d(\xi-x) \frac{H_1^{(1)}(\beta)}{\beta}. \quad (10b)$$

These expressions were evaluated by numerical integration, again using the methods of Ref. [18] and assuming $2w = 10\lambda$, $z = 0.01\lambda$, λ , and 10λ . The results obtained

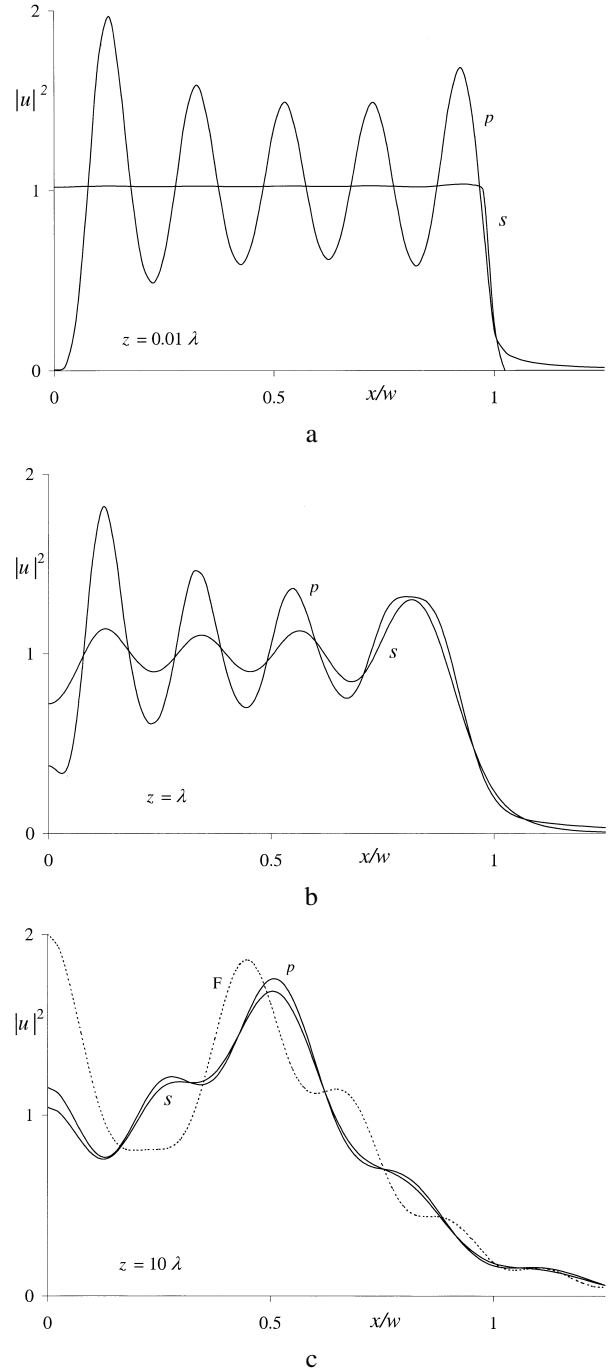


Fig. 5. Diffraction profiles for a circular aperture of diameter $2w = 10\lambda$ at distances z ranging from 0.01λ to 10λ . The profiles shown in each graph are the squared magnitudes of the Rayleigh-Sommerfeld solutions $U_{\text{RS}}^{(p)}$ and $U_{\text{RS}}^{(s)}$ (labeled p and s), and in addition the dotted line in Fig. 5c represents the Fresnel approximation u_F (labeled F).

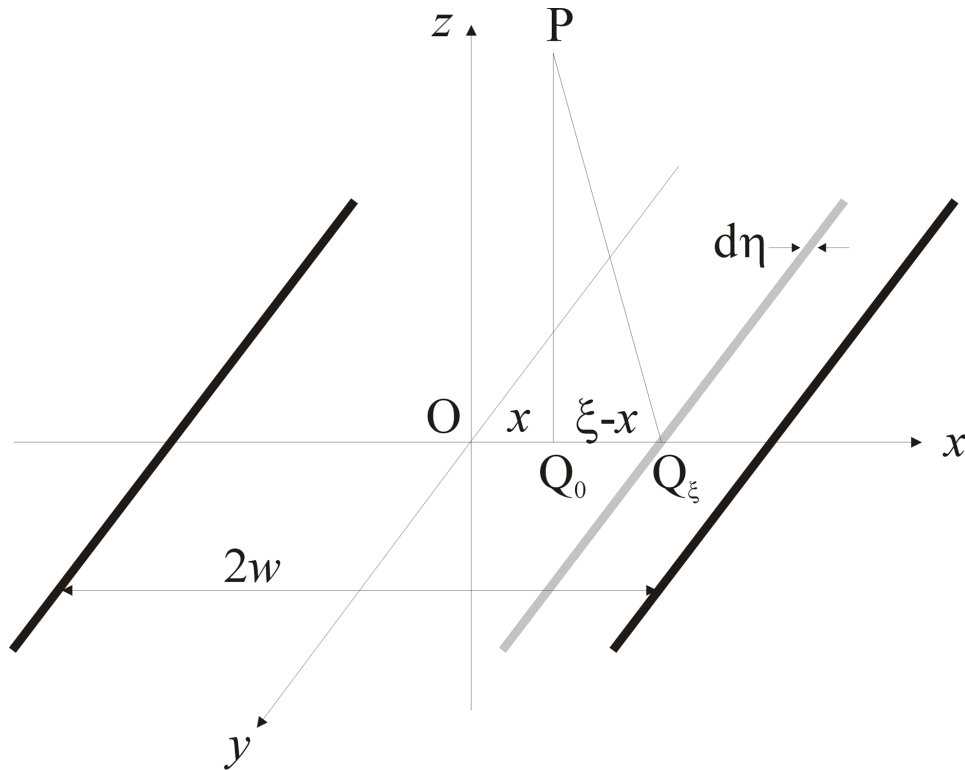


Fig. 6. Geometrical notation used for slits.

are shown in Figs. 7a through c. It should be noted that, in spite of the singularities of $H_0^{(1)}(\beta)$ and $H_1^{(1)}(\beta)/\beta$ at $\beta = 0$, the computations for $z = 0.01\lambda$ presented no problems as long as sufficiently small summation elements $[\Delta(\xi - x) = 0.01w]$ were used.

4. Discussion

The mathematical expressions derived in the previous Section proved their worth for practical applications, in that the computation of the diffraction profiles plotted in Figs. 5 and 7 posed no problems. The results obtained were everywhere finite, free from singularities, and provided new insight into the nature of diffraction in the close proximity of apertures. In spite of the obvious differences between the profiles pertaining to circular apertures slits, their over-all behavior in the near zone is similar so that it may be conjectured that the following observations are not restricted to these specific aperture forms.²

² Although Kirchhoff's solution is not explicitly mentioned here, its behavior can easily be deduced as it is the arithmetic mean of the Rayleigh-Sommerfeld integrals.

(1) As was to be expected, $u_{RS}^{(s)}$ replicates the assumed rectangle functions Eqs. (6a,b) in the limit $z \rightarrow 0$. However, it does not constitute an analytical continuation of the incident field into the half space $z > 0$ because otherwise $\partial u_{RS}^{(p)}/\partial z$, and thus $u_{RS}^{(p)}$, would also replicate their corresponding boundary values. On the other hand, the discontinuities of $u_{RS}^{(p)}$ do not manifest themselves in the form of sudden jumps, as might be surmised from the “*saltus*” interpretation of Kirchhoff's theory. Instead, they are oscillatory in nature and reminiscent of the manner in which the rectangle functions Eqs. (6a,b) might be approximated by a Fourier series.

(2) When z increases, $u_{RS}^{(p)}$ and $u_{RS}^{(s)}$ gradually converge to the Fresnel's integral in Eq. (3). For the aperture width 10λ assumed in the examples the Fresnel limit is expected to be reached when $z \sim 100\lambda$, and yet Figs. 5c and 7c show that the differences between $u_{RS}^{(p)}$ and $u_{RS}^{(s)}$ are already very small at only one tenth this distance. This suggests that, according to the Rayleigh-Sommerfeld theory, polarization effects are negligibly small even in the near zone.

Although the mathematical expressions derived in this paper will be useful for computations in the near zone, it remains unclear which of them ought to be used in given cases. In a pragmatic sense this may be a mute

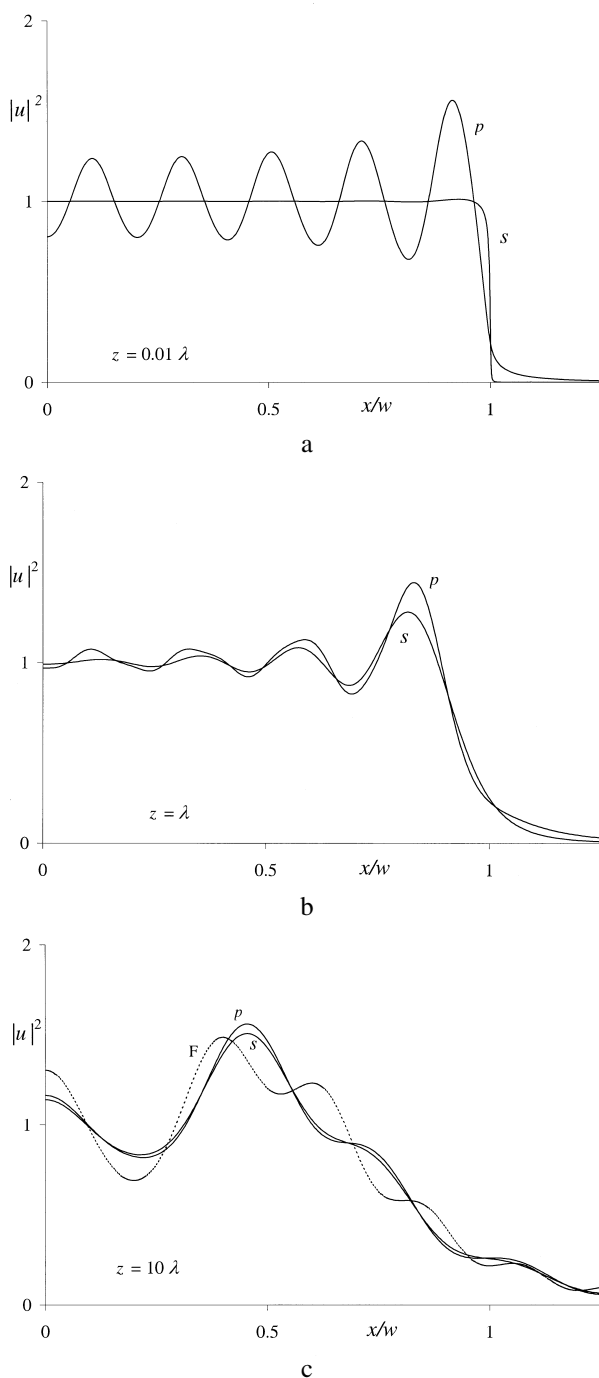


Fig. 7. Diffraction profiles for a slit of width $2w = 10\lambda$ at distances z ranging from 0.01λ to 10λ . The profiles shown in each graph are the squared magnitudes of the Rayleigh-Sommerfeld solutions $U_{RS}^{(p)}$ and $U_{RS}^{(s)}$ (labeled p and s), and in addition the dotted line in Fig. 7c represents the Fresnel approximation u_F (labeled F).

question, because $u_{RS}^{(p)}$ and $u_{RS}^{(s)}$ are so similar in most of the near zone that either will be an improvement over the Fresnel approximation and it does not matter which is used. Therefore, $u_{RS}^{(s)}$ could be (and has been) regarded as the preferred solution as it is continuous in the aperture plane, or Kirchhoff's solution u_K could be regarded as a best compromise as it is the arithmetic mean of $u_{RS}^{(p)}$ and $u_{RS}^{(s)}$. In this author's opinion, these are unfounded guesses. The fact of the matter is that assessing the physical significance of the Rayleigh-Sommerfeld integrals requires additional considerations that will be the subject of a subsequent paper.

5. References

- [1] G. R. Kirchhoff, *Ann. Phys.* **18**, 663 (1883).
- [2] Lord Rayleigh, *Phil. Mag.* **43**, 259 (1897).
- [3] Lord Rayleigh, *Proc. Roy. Soc. (A)* **89**, 194 (1913).
- [4] A. Sommerfeld, *Optik*, Dieterich'sche Verlagsbuchhandlung, Wiesbaden (1950) pp. 200-205.
- [5] H. L. F. v. Helmholtz, *J. Mathematik* **57**, 7 (1859).
- [6] H. Poincaré, *Théorie Mathématique de la Lumière*, G. Carré, Paris 1892 pp. 185-188.
- [7] A. Sommerfeld, *Nachr. Kgl. Akad. Wiss. Göttingen* **4**, 339 (1894).
- [8] N. Mukunda, *J. Opt. Soc. Am.* **52**, 336 (1962).
- [9] K. D. Mielenz, *J. Res. Natl. Inst. Stand. Technol.* **103**, 497 (1998).
- [10] E. Wolf and E. W. Marchand, *J. Opt. Soc. Am.* **54**, 587 (1964).
- [11] A. Rubinowicz, *Ann. Phys.* **53**, 257 (1917).
- [12] H. Osterberg and L. W. Smith, *J. Opt. Soc. Am.* **51**, 1050 (1961).
- [13] F. Kottler, *Ann. Phys.* **70**, 405 (1923).
- [14] E. W. Marchand and E. Wolf, *J. Opt. Soc. Am.* **52**, 761 (1962).
- [15] E. W. Marchand and E. Wolf, *J. Opt. Soc. Am.* **56**, 1712 (1966).
- [16] M. Born, *Optik*, Springer Verlag, Berlin (1930) p. 151.
- [17] W. Franz, *Z. Phys.* **126**, 563 (1949).
- [18] K. D. Mielenz, *J. Res. Natl. Inst. Stand. Technol.* **105**, 581 (2000).
- [19] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Acad. Press, New York (1980) p. 957.
- [20] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, U.S. Gov. Printing Office (1972) p. 358.

About the author: Klaus D. Mielenz is a physicist and retired Chief of the Radiometric Physics Division of NIST Physics Laboratory. The National Institute of Standards and Technology is an agency of the Technology Administration, U.S. Department of Commerce.